A Binocular LVA Device based on Mixed Reality to Enhance Face Recognition

Josselin Gautier, Nicolas Chleq, Pierre Kornprobst

A Binocular LVA Device based on Mixed Reality to Enhance Face Recognition

Josselin Gautier, Nicolas Chleq, Pierre Kornprobst

RESEARCH REPORT
N° 9216
October 2018
Project-Team Biovision

ISSN 0249-6399 ISRN INRIA/RR--9216--FR+ENG

A Binocular LVA Device based on Mixed Reality to Enhance Face Recognition

Josselin Gautier*, Nicolas Chleq*, Pierre Kornprobst

Project-Team Biovision

Research Report n° 9216 — October 2018 — 16 pages

Abstract: Today's visual enhancement systems for low vision people consist of dedicated augmented reality hardware allowing to magnify or enhance the overall scene, independently of the image content or patient needs. For example, for patients with central vision loss, interacting with other has become a painful activity since faces and expressions can hardly be recognized. We introduce a new augmented reality system allowing to selectively enhances faces, using two image processing techniques. Our system has the capacity to adjust the enhancement to the detected faces' size and distance, hence maintaining a constant boost in the critical range of spatial frequency. It offers a binocular and large Field-of-View and performs at near real-time with a modest laptop computer using multithreading. Preliminary experiments with three patients with central vision loss suggest that the enhancements chosen strongly depends on each patient's condition and lead to improved recognition abilities when patients find their optimal settings.

Key-words: low vision, visual aid, sensory aid, image enhancement, head mounted display, augmented reality, mixed reality

* Université Côte d'Azur, Inria, France

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Un dispositif binoculaire LVA basé sur la réalité mixte pour améliorer la reconnaissance des visages

Résumé : Les systèmes d'amélioration visuelle actuels pour les malvoyants consistent en un matériel dédié à la réalité augmentée permettant d'agrandir ou d'améliorer la scène dans son ensemble, indépendamment du contenu de l'image ou des besoins du patient. Par exemple, pour les patients présentant une perte de vision centrale, l'interaction avec d'autres patients est devenue une activité douloureuse, car les visages et les expressions sont difficilement reconnaissables. Nous introduisons un nouveau système de réalité augmentée permettant d'améliorer de manière sélective les visages, en utilisant deux techniques de traitement d'images. Notre système a la capacité d'ajuster l'amélioration à la taille et à la distance des visages détectés, maintenant ainsi une augmentation constante de la plage critique de fréquences spatiales. Il offre un grand champ de vision binoculaire et fonctionne proche du temps réel avec un ordinateur portable modeste utilisant la technologie multithreading. Des expériences préliminaires menées auprès de trois patients présentant une perte de vision centrale suggèrent que les améliorations choisies dépendent fortement de l'état de chaque patient et permettent d'améliorer les capacités de reconnaissance lorsque les patients trouvent leurs paramètres optimaux.

Mots-clés : basse vision, aide visuelle, aide sensorielle, amélioration des images, système de visualisation monté sur la tête, réalité augmentée, réalité mixte

1 Introduction

The growth and change in age structure of the worldwide population is causing a growth in the number of visually impaired people (prevalence) which appear to be accelerating [2]. The number of people with moderate and severe vision impairement is expected to reach 237 millions by 2020 and to triple to 587.6 millions by 2050 [2]. Of great concerns are the people with age-related macular degeneration (AMD), the leading cause of vision loss and blindess among Americans aged of 65 years. Stargardt disease, also called Juvenile Macular Degeneration (JMD), is another leading cause of central vision loss, a macular distrophy mainly caused by a mutation of the ABCA4 gene. People with Stargadt disease experience gray, black or hazy spots in their central vision. Hence low-vision condition afflicts people in their basic daily activities requiring central vision, such as reading [5, 32], recognizing faces [25] or watching television.

There is currently no effective medical treatment reversing the central vision loss associated with AMD nor Stargardt disease (but see [6]). Visual aids such as magnifiers have been used for centuries but remain optically constraint by their magnifying power, field of view (FoV) or viewing distance. Nowadays commonly prescribed Low-Vision Aids (LVA) also include over-correction with plus lenses, optical or digital stand magnifiers and closed circuit television systems (CCTV). They improve reading ability by including brightness and contrast control, color inversion etc. These might be an intermediate step in the evolution of LVA toward image-processing and artificial-intelligence portable systems.

Over the past two decades several publications and commercial applications have proposed the use of head-mounted display technologies (HMD) to enhance human vision (see [37, 26, 17] for reviews on LVAs and [7] for a review on LVA based on HMD). Despite a second wave of commercially available HMD LVA (see table I in [7]), few research has been dedicated to the development and evaluation of HMD LVA i) in practical scenarios and ii) for enhancement of functions other than text reading. Most research on image processing techniques for low-vision enhancement have been proposed and evaluated on static screen, but with a significant number of patient and control.

We propose to go beyond the limitation of existing studies, prototypes and devices by developing and evaluating a portable, binocular, near real-time, high resolution HMD LVA focused on locally enhancing face recognition. The motivation is that for patients with central vision loss, interacting with other has become a painful activity since faces and expressions can hardly be recognized.

The paper is structured as follows. In Sect. 2, we discuss what are the main image processing techniques used in LVA. In Sect. 3, we present our prototypes based on HMD. In Sect. 4, we focus on the algorithms chosen for enhancement and implementation details. Finally, we provide in Sect. 5 our results from a pilot experimentation with patients.

2 Overview of Image Processing Techniques

Patients with low-vision experience a reduction of visual acuity and a loss of contrast sensitivity. Central vision loss add to this combination a central scotoma where the vision is highly reduced or non-existent. Figure 1 illustrates the three different classes of image processing used in LVA: contrast, spatial and edge-based enhancements, scene retargeting and movement-based methods.

While there exist a variety of image processing for low-vision [24], we focus here on describing the most studied ones performing *image enhancement* rather than simplification or motion generation. However as we will see, some techniques in the spatial frequency domain bring a form of scene simplification.

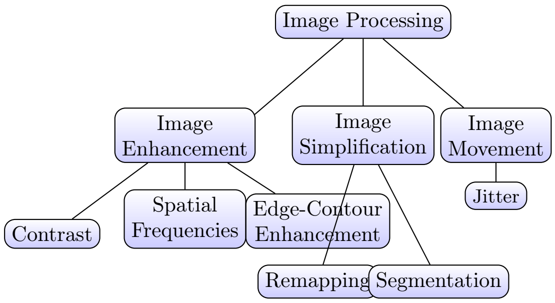

Figure 1: A classification of image processing techniques for LVA

Contrast Contrast enhancement is a standard technique in the sense of its low computational complexity and hence relative long history. On an extended study comparing subsequent methods, Leat et al. [21] found that low-vision patients (N=24) preferred images enhanced by contrast-stretching (with slope and X-intercept varying from (1.5,43) to (3,85) respectively) when tested on CRT screen at 50cm distance. Patients exhibited increase performance at recognizing facial expressions. Peli's team evaluated contrast enhancement on images in the Discrete Cosinus Transform domain on JPEG images ([23, 33]) and MPEG videos ([10, 18]). Low-vision patients exhibited a preference over original images but no improvements during object visual search.

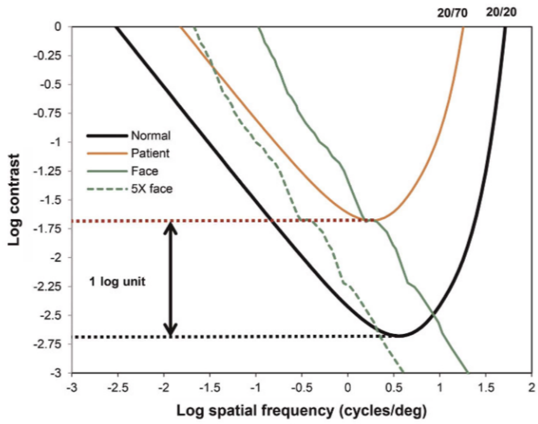

Figure 2: Functions of log contrast threshold versus log spatial frequency for a normally sighed (black) and a low vision patient (orange curve). Reproduced from [7].

Spatial Frequencies In the spatial frequency domain (see Figure 2), a reduction in visual acuity can be interpreted as a reduction of the contrast threshold function toward lower spatial frequency (along the X axis), while a reduced contrast sensitivity will reduce the function amplitude. As illustrated, a magnification or overall contrast increase will not suffice to enable patients to perceive all the spatial frequencies making a face when normally experienced (from 1 to 10m).

In their seminal paper from 1984 [31], Peli and Peli proposed to employ an adaptive filtering method -originally proposed ref() for general image enhancement- which increase the gain of high spatial frequencies while limiting the low-spatial frequencies (to allow a greater dynamic range). Consecutive studies on the visually impaired with text, images and video have shown improved face recognition on static images or better recognition of detail in video, and a general preference over the original. However, one main question regarding the effective range of spatial frequency (in cycle/face) where the boost should occur remains unresolved [29, 30]. In this regard, Lawton proposed to use compensation filters adjusted on patient's normalized contrast sensitivity function and found improved reading performance [20]. Fine and Peli [9] could not replicate this result either because the enhancement parameters were not adjusted to the patient contrast sensitivity, or because of the small sample size in Lawton [20]. Finally, Everingham et al. [8] proposed an image segmentation method targeting mobility based on object color coding. 16 visually impaired patients improved their obstacle recognition by 100% while Peli adaptive enhancement reduced significantly performance compared to original object images. Hence one might need to be careful as image enhancement needs to be adjusted to patient's visual disorder, visual acuity, and visual task.

Edge and contour Edge-based enhancement is appealing as it has been known for decades that the primary visual cortex (V1) is performing edge detection [15]. In its simpler form where the contrast of edges is boosted [19], an edge enhancement is apparent to high pass filtering in the spatial frequency domain. Superimposing colored edges, despite little evaluation [1], might disturb the patient on top of distorting the image content. However processing reversed bright edges allow to superimposed them on augmented reality see-through display without obscuring the natural view. Several studies on limited number of patients (2 [22], 3 [16], 8 [28]) have shown improved contrast sensitivity [22, 16], visual acuity ([28, 22]) and increased visual field [28].

2.1 HMD-based Low Vision Aids

Head Mounted Displays (HMD) offer among today's best solution to display non-invasively processed images onto the eyes. However, past and present solutions still suffer from weight, bulkiness, reduced peripheral fields, and of course impossibility of eye contact. While to our knowledge only two prototypes of image processing contour-based technique have been proposed and tested with See-Through Displays (STD) bringing *vision multiplexing* [28, 27, 16], none have to date evaluated Image Processing directly into an HMD. However, while STD suffer from major limitations in term of dynamic range, contrast and especially reduced field of view (which makes it impractical for disorders other than peripheral vision defects loss such as Glaucoma, Retinitis Pigmentosa), HMD carries important potential with large field of view, constant illumination and high angular resolution which could reveal practical for people with central vision loss. Few studies have evaluated on patients HMD LVA, their usability nor their performance. Other than commercial devices performing mainly digital magnification (eSight, NuEyes,etc.), HMDs have been tested in research for text enhancement [12] or scotoma simulation on normal subjects. The other HMD-based devices intending LVA focused on the engineering aspect and its practical challenges: real-time (GPU) [34] or FPGA-based, [3, 11], user-configurability [11]. Interestingly this last method aims to let the user switch in real-time between edge gradient techniques, their magnitude and thickness. In conclusion several groups have studied over the past two decades separate aspects of LVA design: either image-processing techniques testing face recognition performance on screen, or practical implementation ensuring real-time performance under constraint of portability (low energy consumption).

3 Proposal

Our motivation is to embed visual enhancement techniques fed from direct stereoscopic camera video stream into a portable device whose resulting frames would be displayed in near real-time to a HMD. We developed two prototypes describe below:

3.1 Prototype 1

The first HMD prototype was developed to test the feasability of embedding a real-time image processing performing 1 (Peli)[31] or 2 (xDoG)[38] convolutions on top of arithmetic operations on every full video frame captured from the front camera of a smartphone (Samsung Galaxy S7, video 720p @30fps), inserted into a Samsung Gear VR headset. We reasoned that processing the whole image without performing live face detection would give a decent estimate of the frame rate and an idea of the rendering abilities and limitations. Using a shader computation of the Peli's method [31], embedded into an Unity Smartphone app, this first prototype shows an overall satisfying rendering, with very strong contrast somehow visually too sharp for the normal subject. The framerate did not exceed 10 fps however, which motivated us to move to a more flexible and powerful platform.

3.2 Prototype 2

Conserving binocular vision for people with low-vision offers several critical advantages: it provides them with binocular depth cues that in spite of their poor central vision, they can employ to better estimate distance along the horopter (up to 2m). It has been shown that coarse stereopsis may benefit tasks of daily living for individuals with central field loss[35]. Helping binocular summation by sensory and motor fusion also keep the vergence system effective. Finally having to read -hence at close distance- prints with potential faces to enhance would elicit near convergence and then require stereoscopic viewing.

To that end, we will first describe the hardware (see Figure 3): the stereoscopic camera acquisition by the LVA, the display and then the software, i.e. the proposed processing.

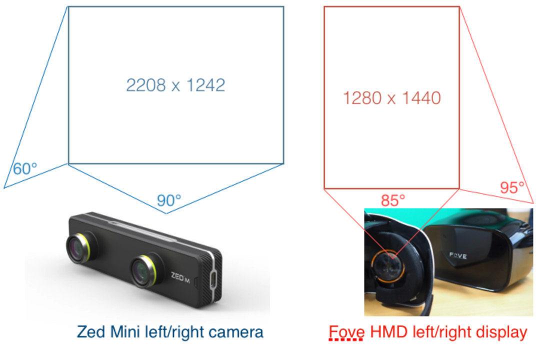

Figure 3: Illustration of the horizontal and vertical fields of view and corresponding native resolutions of each camera and each embedded display.

Cameras Despite the move toward augmented reality, existing HMDs to date do not come with more than one camera (but see See-through *Hololens* or unreleased *VrVana*). Hence we propose to place a dedicated stereo-camera rig in front the corresponding optic primary axis of each eye. The *ZED mini* stereo camera (Stereolabs Inc., San Francisco, CA, USA) is dedicated to mixed reality. It captures HD videos with a naturalistic cameras' Inter Pupillary Distance (IPD or stereo baseline) of 63mm and weights 63 grams. Its field of view is 90° (H) x 60° (V). In practice, its field of view changes with the input resolution chosen: a chart test measurement indicate actual 95° (H) x 56.3° (V) @ 720p per view. A first direct input-to-output test make us suffer from diplopia. We realized with printed chart that the two cameras had their two optical axis misaligned of few vertical degrees. While the Human Visual System (HVS) copes daily with horizontal misalignment of its visual axis, it is very sensitive to any vertical deviation (or vertical disparity) larger than few minutes of arc. Because in euclidean geometry a rotation in 3D is equivalent to a translation in the 2D projected space, we correct this vertical misalignment by manually translating vertically the right view. A chart test with a regular grid validated our subjective test. In practice, because the whole right image has been translated upward, some top horizontal video bands on the right view and bottom band on the left view now needs to be trimmed to avoid unpleaseant vertical shift, which generate an extra loss of vertical field of view.

Head Mounted Displays We chose to the use of a large field of view (FoV) HMD that do not sacrifice the angular resolution, using a Fove 0 HMD (FOVE Inc, San Mateo, CA, USA). Its monocular FoV (85° (H) x 95° (V)) is higher and thinner than each corresponding frontal camera FoV, while its resolution per view (1280(H) x 1440(V) pixels) do not correspond neither to the native resolution of the camera (see 3). As we will see later, we chose to use a video acquisition mode of the Zed-mini camera to provide the best visual experience by trading-off latency on resolution. We also chose this HMD that embeds a video eye-tracker to further develop gaze-contingency features.

4 Implementation

While the first prototype was initially developed in C# with Unity 3D, we switched to a faster C++ implementation using OpenCV image processing library. Our general approach was to multi-thread (on the CPU) the two operations of face detection and acquisition-face enhancement-display separately (see subsection 4.2).

4.1 Face Detection

To devote the CPU and potential GPU resources to face image processing, we decided to use a simple Haar-based feature cascade classifier from OpenCV [36]. Specifically, for face detection the `haarcascade_frontal face_alt.xml` classifier was used. In order to enhance faces within a acceptable range of distance from 1m to 5m, we restrict the detection to small and large limits (40 pixels and 180 pixels respectively for a single view in 1280x720 pixels). Also, without gaussian cascaded implementation, the convolution processing time increase exponentially with the face size, so it is critical to put an upper limited to the detected face size. Because this classifier is multiscale, we need to specify how much an image is resized from one scale to another. Here it is set to 1.1.

A specific contribution consist in detecting the face on one view only. Because the distance to a detected face his proportional to its size, and because the stereo disparity of that face between the two view is also proportional to its distance, then it comes that we can estimate the disparity of the face in view 2 from its position and **size** in view 1, as illustrated in Figure 4.

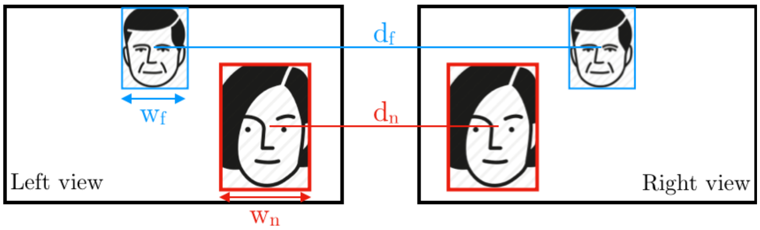

Figure 4: Principle for the monocular face detection on left view and face repositioning on right view

4.2 Face Tracking and Multithreading

On the target hardware, the face detection process is fast but not sufficient to cope with the input framerate. It also exhibits both spurious and failed detections, both at a very low rate. Without further processing this produces a flickering effect in the output. The multithreading of tasks was arranged to overcome these limits:

- One thread repeatedly runs the face detection on the most recent acquired frame and produce raw detections in a shared memory buffer. The detection rate is lower than the input frame rate, and the ratio between these rates is in the range (2,5). This ratio was found to give visually acceptable results.
- Another thread tracks these faces to overcome spurious and failed detections. The tracking process matches successive raw detections with simple matching criterions based on geometrical computations. It also uses a weighting mechanism to reflect both the continuity of detections and to recover from transient detection failure. It produces a set of stable detections used by the enhancement tasks on every input frame, with a noticeable reduction in the flickering effect.

4.3 Image Enhancement Techniques: Peli and xDoG

In this paper we have focused on two image processing techniques. The first is the Peli transform [29] which has been proposed for low vision people. Given the image I to enhance, the equation of the Peli's adaptive enhancement is as follow:

$$P_\sigma(\hat{x}, \sigma, k, n) = k \cdot I_{high-pass}(\hat{x}) + n \cdot I_{low-pass}(\hat{x}) + I, \tag{1}$$

with

$$I_{high-pass}(\hat{x}) = I(x) - G(\hat{x}, \sigma) \quad \text{and} \quad I_{low-pass}(x) = G(x, \sigma) - I,$$

where σ, k, n are parameters and

$$G(\hat{x}, \sigma) = \frac{1}{2\pi\sigma^2} \int I(x) \mathrm{e}^{-\dfrac{\|\hat{x} - x\|}{2\sigma^2}} \mathrm{d}x.$$

Hence Peli's adaptive enhancement can be summarised as a high-frequency spatial filter of gain gain k and frequency cutoff σ whose gain n in the low spatial frequency band is controlled. This was proposed to avoid distorting suprathreshold low-SF features while maintaining the display range for enhancement of higher frequency subthreshold features.

The second approach we considered comes from the computer graphics domain for image stylization. It is called the xDoG transform, proposed by Winemöller et al. [38]. Because of its qualitative effect on faces and its computational performance, it could also be a good candidate in the context of LVA. It can be considered as an extension of the classical Difference of Gaussian (DoG) classically used in image enhancement. More precisely, given an image I, it xDoG transform is defined by:

$$E(\sigma, k, \tau, \epsilon, \varphi) = \begin{cases} 1 & \text{if } D(\sigma, k, \tau) < \epsilon \\ 1 + \tanh(\varphi \cdot (D(\sigma, k, \tau))) & \text{otherwise} \end{cases} \tag{2}$$

where

$$D(\sigma, k, \tau) = G(\sigma) - \tau \cdot G(k \cdot \sigma).$$

5 Results

The results cover both the development aspect by reporting the performance of the software face detection, and the subsequent results from the patients' point of view, in a realistic live scenario of face recognition in a testing room simulating a flat environment.

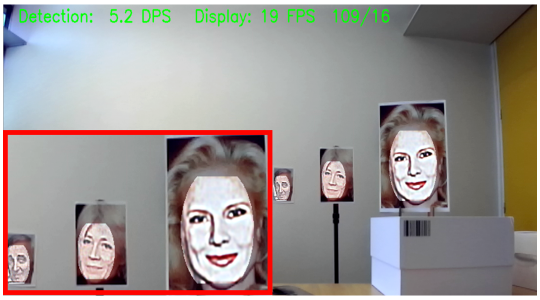

Figure 5: Rendering of xDoG enhancement on the left view for 3 french celebrities faces at 1, 2 and 3m from the large FoV camera. The adjustment of spatial frequency cutoff to the distance of faces can be appreciated into the inset (in red)

5.1 Algorithm performance

The multithreaded software was tested with a modest 14" laptop from 2014 (HP EliteBook i7@2.1 GHz), that also served during the evaluation. We report in Figure 6 the performance of the algorithm in term of detection of face per second (on the first thread), and the output frame per second. The multithread does not guarantee that a detected face will be processed for every frame. At the opposite, we observed that face can be detected up to 5 times the acquisition framerate without consequences. However the bottleneck limiting real-time ability (30 fps) is at the image processing with costly convolution. This saturates the end-to-end framerate to 10 and 15 fps for Peli and xDoG respectively.

5.2 Experiments with Low-Vision Patients

The goal of this pilot experimentation is fourfold: evaluate the general comfort and usability in wearing a HMD-based LVA, identify the preferred enhancements, quantify the face recognition performance and estimate the cognitive and physical workload.

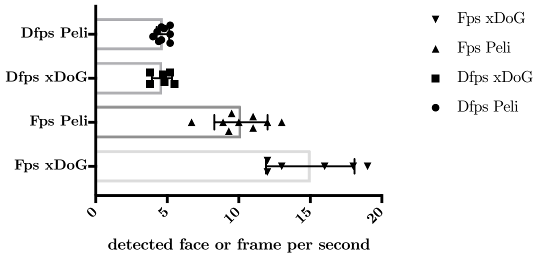

Figure 6: Detected face per second (Dfps) and output Frame per second (Fps)

Patient	Age	Disorder	Visual Acuity (OD/OS)
1	89	AMD	5/10, 4/10
2	71	Stargardt	1/10, 1.25/10
3	92	AMD	0.8/10, 2/10

Table 1: Patients' age, condition and visual acuity

Patients Patients were selected by the opthalmologic service from the Pasteur 2 university hospital of Nice (CHU), based on their motivation to test different research prototypes and the apparent wide range of technology acceptance they represent 1. The experiment was conform to the Declaration of Helsinki and written informed consent was obtained from each subject before the experiment.

Face Recognition performance evaluation The patient's ability to recognize celebrities was tested to evaluate his face recognition performance, rather than his ability to discriminate among a limited selection of test faces (i.e. a matching task). It has also been argued that familiar face recognition is more robust than the recognition of recently memorized faces (ref(seeinPeli)). While it would have been interesting to include a subset of familiar faces from each of the patient's relative, a consequent work would have been necessary to have avoid memory bias with 'iconic' pictures. 34 faces of french (singers, politician) and international (politicians) celebrities were selected based on their notoriety covering the past to present age of patients.

Cognitive and physical effort evaluation The NASA-TLX [14] was developed by Human Performance Group at Ames Research Center as a subjective multidimensional assessment tool where subject are asked to rate their perceived workload to achieve a task. It is now a widely used (4400+ studies [13] in a variety of domains. The approved french translation of the NASA-TLX was used [4].

Protocol The protocol was devised to favor a progressive handling of the head mounted low-vision device of significant weight (520+63=630 grams without cables). First, 6 printed faces were shown to the patient without any device, at 1.2m. Second, after a slow adjustment of the LVA to the patient head and glasses (if present), he was proposed to subjectively report his preferred *"face display"* between Peli xDoG at three different levels of gain. Then the headset was removed for a break. The subject was introduce to the handling of a video game controller

Method	Parameter	variable	range	Simp. name
	HF gain	k	[1:1:10]	visibility
Peli	HF F_{cutoff}	σ	[1:1:14]	hardened
	n	SF gain	[0.5:0.5:1]	(contrast)
	Gain	p	[20:10:80]	luminosity
xDoG	HF & LF	σ	[0.1:0.1:1 ...	visibility
	F_{cutoff}		... 1:1:6]	
	contrast	φ	[10:10:80]	(contrast)

Table 2: Parameters and corresponding range for the two proposed method, and their enunciated name

and four of its buttons, to adjust both gain and frequency of the previously selected method within a predefined range (see table 2). The whole protocol is summarized below.

Step1 Celebrity face recognition without LVA, at 1.2m

Step2 Handling of LVA and selection of preferred enhancement for 3 values of gain each (3*2).

Step3 Self-adjustment of gain and frequency (see table 2). Face recognition evaluation.

Step4 Cognitive and physical effort assessment (NASA-TLX)

5.3 Subjective report and preferences

The step 2 of the experiment allowed to identify the preferred face enhancements, if there were. Patient 1 reported a preference for the xDoG method, while patient 2 preferred the Peli's one. Patient 1 first exclamate when adjusting the headset:
≪ *Oh, I can see you in 3D ! (...) your face has more relief, I can better see you eyes, their shadows, etc...* ≫. He was appreciating both the luminance (p) and contrast boost (parameter φ) as he latter push them to nearly their maximum ($p = 50, \varphi = 70$).
Patient 2 was at first glance more circumspect by the rendering of the faces, that he reported less natural. However as we will see in section 5.4, he dramatically improved his face recognition ability. Interestingly patient 2 also suggest at the end of step 3 to combine an increase in contrast with a magnification of the face (a zoom-in feature).

This step 2 also revealed that it was difficult to embed Patient 3's spectacles into the *Fove* HMD. These trifocals spectacles fit hardly into the front plastic part of the HMD. At this step, patient 3 reported not perceiving any visual difference while the contrast was increased, the spectacles removed or not, and the distance to face varied. Consequently results are reported for Patient 1 and 2, while the limitation for patient 3 are discussed in section 6.

5.4 Recognition performance

As illustrated in Figure 7 (left), Patient 1 shown reduce recognition performance at 1.2m while being positive about his perception. At the opposite, patient 2 who was relatively circumspect shows a large and systematic improvement in recognizing faces. Because he was not able to recognize any faces at 1.1m distance without the headset, we decided to test approx half the distance by bringing the printed faces to be recognized to 0.5m. In both cases, the LVA helped greatly the patient to recognize faces, from no recognition to 50% and 83% at 1.1 and 0.5m respectively.

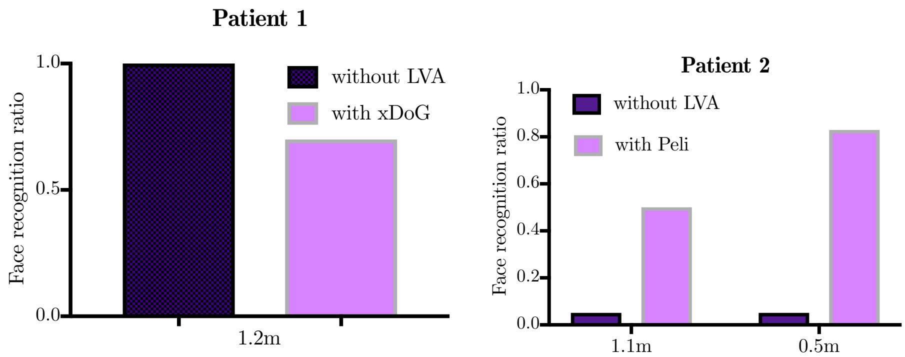

Figure 7: Recognition performance with the proposed LVA for patient 1 (left) and 2 (right) on faces of past and present celebrities.

5.5 Subjective assessment

The results of the NASA-TLX present a great variability between the two patients. Patient 1 appear moderately affected by the task except in term of mental demand, and recognize middle-to-low performance. For patient 2, he observed having devoted significant effort to a subjectively reported physically and mentally demanding task, but on the other hand also recognized having very well performed to recognize face.

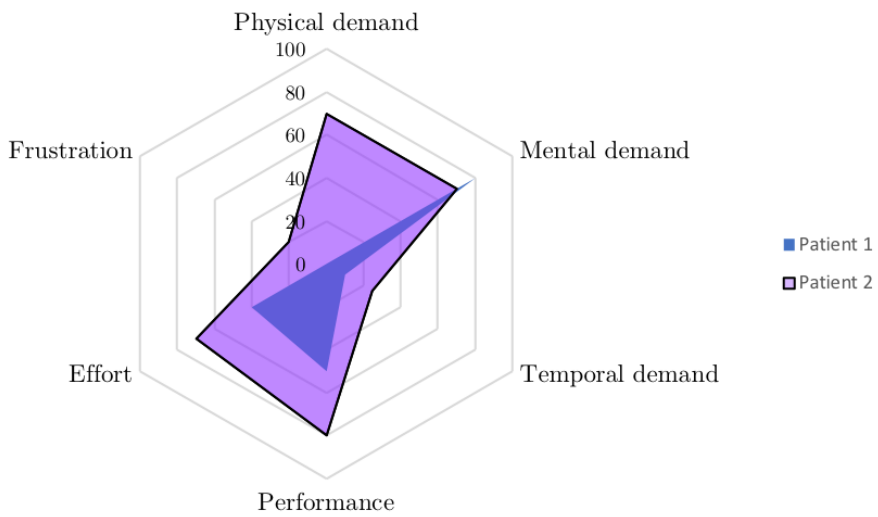

Figure 8: Result of the NASA task load index

6 Discussion

Image enhancement can provide a measurable improvement in face recognition for the visually impaired people. However, for any assistive device evaluation it remains critical to transpose reported subjective preference or objective gains to realistic case scenario. Hence the double motivation to first develop a binocular large FoV device for low-vision people that would not impede their peripheral vision nor their even low stereo vision, whose objective would be to supplement the critical task of face recognition. Patients' face recognition performances go *a priori* again their subjective reports of confidence and usability, which shows the necessity not to stop to qualitative feedbacks but to objectivize the measurements and take into account the patient's psychology and acceptability to new-technology, especially at this older age. While patient 1 appeared very positive (notes from video recordings), he exhibited lower face recognition (but on a small sample of faces, limited by the conditions of the experiment). Patient 2, while quite critical, improved dramatically his face recognition ability. Patient 3 also provides an interesting case for developing adapted HMD (e.g. with adaptive focal plane). The subject assessment of cognitive and physical workload also gives precious insights on the across-subject variability when learning to use a new technology. Age, patient's condition, motivation, technology acceptability, all conditioned the success in low-vision aid adoption among elderly population.

Figure 9: Illustration of the experimentation in a realistic environment. A patient is wearing a headset while being asked to judge a portrait at intermediate distance

7 Conclusion

A new kind of binocular low-vision assistive device aiming at facilitate face recognition has been developed and evaluated in real conditions. Several contributions have been proposed and tested that proved today's maturity and convergence of portable technologies to prototype, develop and evaluate realistic visual aids for people with impaired vision. From the engineering point of view, the feasability of live face enhancement has been shown with practical contribution to speed up the process. The LVA system can adjust its enhancement to the detected face'size and distance while being also controllable by the user. From the visual enhancement and low-vision perspective, several questions have been open, inviting to continue the research by first relating on-screen patients' visual preferences (gain and frequency) in agreement with their contrast sensitivity across spatial frequencies, and second by evaluating practical performances on a daily basis.

Acknowledgments

This work was partially supported by Bosch Visiontec in Sophia Antipolis, Valbonne, France. We would like to thank David Coupé and Frédéric Dosière from Bosch Visiontec for their technical and financial support, Prof. Stéphanie Baillif from CHU Pasteur 2 for allowing us to meet patients with central vision diseases, David Rey and Guilene Ribiere for their support throughout this short project, and Marco Benzi for technical support.

References

[1] W Atabany and P Degenaar. A robust edge enhancement approach for low vision patients using scene simplification. In *Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International*, pages 1–4. IEEE, 2008.

[2] Rupert RA Bourne, Seth R Flaxman, Tasanee Braithwaite, Maria V Cicinelli, Aditi Das, Jost B Jonas, Jill Keeffe, John H Kempen, Janet Leasher, Hans Limburg, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *The Lancet Global Health*, 5(9):e888–e897, 2017.

[3] P Martínez Cañada, C Morillas, Raquel Ureña, JM Gómez López, and Francisco J Pelayo. Embedded system for contrast enhancement in low-vision. *Journal of Systems Architecture*, 59(1):30–38, 2013.

[4] Julien Cegarra and Nicolas Morgado. Étude des propriétés de la version francophone du nasatlx. In *Communication présentée à la cinquième édition du colloque de psychologie ergonomique (Epique)*, pages 28–30, 2009.

[5] Susana TL Chung, J Stephen Mansfield, and Gordon E Legge. Psychophysics of reading. xviii. the effect of print size on reading speed in normal peripheral vision. *Vision research*, 38(19):2949–2962, 1998.

[6] Lyndon da Cruz, Kate Fynes, Odysseas Georgiadis, Julie Kerby, Yvonne H Luo, Ahmad Ahmado, Amanda Vernon, Julie T Daniels, Britta Nommiste, Shazeen M Hasan, et al. Phase 1 clinical study of an embryonic stem cell–derived retinal pigment epithelium patch in age-related macular degeneration. *Nature biotechnology*, 36(4):328, 2018.

[7] Ashley D Deemer, Christopher K Bradley, Nicole C Ross, Danielle M Natale, Rath Itthipanichpong, Frank S Werblin, and Robert W Massof. Low vision enhancement with head-mounted video display systems: Are we there yet? *Optometry and Vision Science*, 95(9):694–703, 2018.

[8] MR Everingham, BT Thomas, T Troscianko, et al. Head-mounted mobility aid for low vision using scene classification techniques. *The International Journal of Virtual Reality*, 3(4):3, 1999.

[9] Elisabeth M Fine and Eli Peli. Enhancement of text for the visually impaired. *JOSA A*, 12(7):1439–1447, 1995.

[10] Matthew Fullerton, Russell L Woods, Fuensanta A Vera-Diaz, and Eli Peli. Measuring perceived video quality of mpeg enhancement by people with impaired vision. *JOSA A*, 24(12):B174–B187, 2007.

[11] Ryan M Gibson, Ali Ahmadinia, Scott G McMeekin, Niall C Strang, and Gordon Morison. A reconfigurable real-time morphological system for augmented vision. *EURASIP Journal on Advances in Signal Processing*, 2013(1):134, 2013.

[12] Anshul Gupta, Juraj Mesik, Stephen A Engel, Rebecca Smith, Mark Schatza, Aurélie Calabrèse, Frederik J Van Kuijk, Arthur G Erdman, and Gordon E Legge. Beneficial effects of spatial remapping for reading with simulated central field loss. *Investigative ophthalmology & visual science*, 59(2):1105–1112, 2018.

[13] Sandra G Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage Publications Sage CA: Los Angeles, CA, 2006.

[14] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[15] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106–154, 1962.

[16] Alex D Hwang and Eli Peli. An augmented-reality edge enhancement application for google glass. *Optometry and vision science: official publication of the American Academy of Optometry*, 91(8):1021, 2014.

[17] Jeffrey W Jutai, J Graham Strong, and Elizabeth Russell-Minda. Effectiveness of assistive technologies for low vision rehabilitation: A systematic review. *Journal of visual impairment & blindness*, 103(4):210, 2009.

[18] Jeonghoon Kim, Avni Vora, and Eli Peli. Mpeg-based image enhancement for the visually impaired. *Optical engineering*, 43(6):1318–1329, 2004.

[19] MiYoung Kwon, Chaithanya Ramachandra, PremNandhini Satgunam, Bartlett W Mel, Eli Peli, and Bosco S Tjan. Contour enhancement benefits older adults with simulated central field loss. *Optometry and vision science: official publication of the American Academy of Optometry*, 89(9):1374, 2012.

[20] Teri B Lawton. Improved word recognition for observers with age-related maculopathies using compensation filters. *Clin Vis Sci*, 3(3):125–135, 1988.

[21] Susan J Leat, Gloria Omoruyi, Andrew Kennedy, and Ed Jernigan. Generic and customised digital image enhancement filters for the visually impaired. *Vision research*, 45(15):1991–2007, 2005.

[22] Gang Luo, Zhengzhou Li, and Eli Peli. Mobile electronic magnification device for people with central vision loss. *Investigative Ophthalmology & Visual Science*, 52(14):386–386, 2011.

[23] Gang Luo, PremNandhini Satgunam, and Eli Peli. Visual search performance of patients with vision impairment: effect of jpeg image enhancement. *Ophthalmic and Physiological Optics*, 32(5):421–428, 2012.

[24] Howard Moshtael, Tariq Aslam, Ian Underwood, and Baljean Dhillon. High tech aids low vision: a review of image processing for the visually impaired. *Translational vision science & technology*, 4(4):6–6, 2015.

[25] Pia Mäkelä, Risto Näsänen, Jyrki Rovamo, and Dean Melmoth. Identification of facial images in peripheral vision. *Vision Research*, 41(5):599 – 610, 2001.

[26] Nhung Xuan Nguyen, Malte Weismann, and Susanne Trauzettel-Klosinski. Improvement of reading speed after providing of low vision aids in patients with age-related macular degeneration. *Acta ophthalmologica*, 87(8):849–853, 2009.

[27] María Dolores Peláez-Coca, Paloma Sobrado-Calvo, and Fernando Vargas-Martín. Optoelectronic aid for patients with severely restricted visual fields in daylight conditions. *Journal of Modern Optics*, 58(19-20):1871–1879, 2011.

[28] María Dolores Peláez-Coca, Fernando Vargas-Martín, Sonia Mota, Javier Díaz, and Eduardo Ros-Vidal. A versatile optoelectronic aid for low vision patients. *Ophthalmic and Physiological Optics*, 29(5):565–572, 2009.

[29] Eli Peli, Robert B Goldstein, George M Young, Clement L Trempe, and Sheldon M Buzney. Image enhancement for the visually impaired. simulations and experimental results. *Investigative Ophthalmology & Visual Science*, 32(8):2337–2350, 1991.

[30] Eli Peli, Estella Lee, Clement L Trempe, and Sheldon Buzney. Image enhancement for the visually impaired: the effects of enhancement on face recognition. *JOSA A*, 11(7):1929–1939, 1994.

[31] Eli Peli and Tamar Peli. Image enhancement for the visually impaired. *Optical engineering*, 23(1):230147, 1984.

[32] Gary S Rubin and Kathleen Turano. Low vision reading with sequential word presentation. *Vision research*, 34(13):1723–1733, 1994.

[33] Jinshan Tang, Jeonghoon Kim, and Eli Peli. Image enhancement in the jpeg domain for people with vision impairment. *IEEE Transactions on Biomedical Engineering*, 51(11):2013–2023, 2004.

[34] Raquel Ureña, Christian Morillas, and Francisco J Pelayo. Real-time bio-inspired contrast enhancement on gpu. *Neurocomputing*, 121:40–52, 2013.

[35] Preeti Verghese, Terence L Tyson, Saeideh Ghahghaei, and Donald C Fletcher. Depth perception and grasp in central field loss. *Investigative ophthalmology & visual science*, 57(3):1476–1487, 2016.

[36] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 511–518, 2001.

[37] Gianni Virgili, Ruthy Acosta, Sharon A Bentley, Giovanni Giacomelli, Claire Allcock, and Jennifer R Evans. Reading aids for adults with low vision. *Cochrane Database of Systematic Reviews*, (4), 2018.

[38] Holger Winnemöller, Jan Eric Kyprianidis, and Sven C Olsen. Xdog: an extended difference-of-gaussians compendium including advanced image stylization. *Computers & Graphics*, 36(6):740–753, 2012.

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399